Institut royal des Sciences naturelles de Belgique

BULLETIN

Tome XXXI, n° 26
Bruxelles, juillet 1955.

Koninklijk Belgisch Instituut voor Natuurwetenschappen

MEDEDELINGEN

Deel XXXI, nr 26
Brussel, Juli 1955.

ELLIPSES ÉQUIPROBABLES ET TAUX D'ÉLOIGNEMENT EN BIOMÉTRIE,

par Elisabeth DEFRISE-GUSSENHOVEN (Bruxelles).

INTRODUCTION.

En biométrie, on est très fréquemment amené à envisager simultanément deux séries de mensurations entre lesquelles existe un lien plus ou moins étroit.

L'exemple le plus ancien et aussi le plus célèbre est celui des tailles des parents et de leurs enfants présentées dans un tableau à double entrée par F. GALTON (1).

Dès 1900, des tableaux de corrélation analogues se rencontrent dans les sciences naturelles pour la description d'espèces (2) et en anthropologie physique pour l'étude de l'homme (3).

Récemment, plusieurs chercheurs de l'Institut royal des Sciences naturelles de Belgique ont publié de tels tableaux (4); chaque fréquence, indiquée chez F. GALTON par un chiffre, est souvent représentée chez eux par un nombre de points équivalent, de

(1) GALTON, F., 1885, p. 248.

(2) DAVENPORT, C. B., 1903, p. 149. Anteroposterior/Dorsoventral diameter in Lower Valve of *Pecten opercularis*.

(3) PEARSON, K. and LEE, A., 1903, p. 399. Statute/envergure, etc.

(4) GLIBERT, M., 1945, p. 104. Diamètre antéropostérieur/diamètre umbonoventral, etc., des Pélécypodes.

GLIBERT, M., 1949, p. 39. Hauteur/diamètre chez *Calliostoma tauromiliare*.

TWIESSELMANN, F., 1951, pp. 7, 10, 26-29.

JOLIVET, P., 1952, p. 5. Longueur de l'élytre/largeur du pronotum, etc., des Chrysomélides.

ADAM, W., 1952, p. 91. Longueur des bras/longueur du manteau, etc., des Céphalopodes.

sorte que les diagrammes prennent l'aspect de véritables nuages de points (scatter diagrams) où les régions de forte densité correspondent à une haute fréquence.

Ainsi, c'est par des nuages de points que F. TWIESSELMANN illustre la liaison stochastique existant entre certaines mensurations de la population de l'Oasis de Kharga (4) : comme la stature et la taille assis; la longueur et la largeur de la tête; l'indice facial et l'indice céphalique; etc...

Signalons que, dans leur livre récent de méthodologie taxonomique, E. MAYR, E. G. LINSLEY et R. L. USINGER recommandent également l'utilisation de tels diagrammes de dispersion (5).

Une façon élégante de condenser de tels tableaux, sans perdre l'essentiel de l'information contenue dans les données primitives, consiste à utiliser des ellipses concentriques d'égale fréquence ou équiprobables, enfermant un nombre de plus en plus élevé de sujets du groupe.

C'est encore à F. GALTON que revient le mérite de cette découverte. Il crut voir des ellipses dans les lignes qui joignent les points d'égale fréquence de ses tableaux (6) et son ami J. DICKSON, consulté, prouva, par le calcul, que c'est bien d'ellipses qu'il s'agit (7).

Malgré un début si heureux, et bien que la théorie de ces ellipses figure souvent en bonne place dans les traités de statistiques (8), nous ne croyons pas que, par la suite, on les ait souvent utilisées en biométrie.

Actuellement, grâce à la diffusion des techniques statistiques et à un outillage approprié (machine à calculer électrique, ellipsographe) mis à leur disposition, quelques laboratoires de l'Institut royal des Sciences naturelles de Belgique et de l'Association pour l'Étude de la Paléontologie et de la Stratigraphie houillères ont pu reprendre ce mode de représentation. Les résultats de certains de leurs travaux viennent de paraître ou sont en préparation (9).

(5) MAYR, E., LINSLEY, E. G. et USINGER, R. L., 1953, p. 137.
(6) GALTON, F., 1885, p. 255.
(7) DICKSON, J. D. H., 1886, p. 63.
(8) Voir par exemple CRAMÈR, H., 1946, p. 288.
(9) TWIESSELMANN, F., 1954;
VANDERCAMMEN, A.;
PASTIELS, A., 1953;
DEFRISE-GUSSENHOVEN, E., 1954;
Signalons aussi un article de NOWAKOWSKI T. K. et PERKAL, J., 1952, p. 28, où la liaison existant entre le poids et la taille de groupes d'écoliers est illustrée par des ellipses.

Notre but est de souligner l'importance des ellipses équiprobables dans les représentations graphiques. Les ellipses schématisent les données initiales de la population sans les masquer complètement. En effet, on déduit du tracé des ellipses équiprobables tous les caractères biométriques habituels (moyennes, déviations standard, coefficient de corrélation, droites de régression).

Nous introduisons une notion nouvelle, le taux d'éloignement, permettant d'évaluer la distance biométrique entre un sujet aberrant et les moyennes d'une population.

Nous résumons les propriétés essentielles des ellipses équiprobables et du taux d'éloignement sous une forme simple, dont le naturaliste puisse directement déduire l'application dans un problème particulier.

Enfin, trois exemples illustrent l'intérêt pratique des ellipses équiprobables et du taux d'éloignement en anthropologie et en zoologie.

1. — Comment les ellipses équiprobables et le taux d'éloignement s'introduisent tout naturellement dans le raisonnement biologique.

Nous choisissons comme données de base 192 étudiants belges de l'Université de Bruxelles, mesurés en 1945 par F. Twiesselmann et étudiés par cet auteur en 1949 (10).

Tableau I.

Paramètres d'une population d'étudiants belges de l'U.L.B., estimés à partir des valeurs correspondantes de l'échantillon des 192 étudiants.

	Moyenne en cm	Dév. st. en cm	Coefficient de variation	Coeff. de corr. entre x et y
x = stature	X = 174,55	$\sigma_x = 6{,}314$	$v_x = \frac{100\,\sigma_x}{X} = 3{,}61$	
y = tibia	Y = 39,43	$\sigma_y = 2{,}259$	$v_y = \frac{100\,\sigma_y}{Y} = 5{,}73$	$\rho = 0{,}783$
$i = \frac{100y}{x}$ = indice	I = 22,58	$\sigma_i = 0{,}828$	$v_i = \frac{100\,\sigma_i}{I} = 3{,}66$	

(10) Twiesselmann, F., 1949.

Ces 192 étudiants constituent l'échantillon représentatif de la population plus vaste de tous les étudiants belges de l'U. L. B. en 1945.

Ne retenant pour chaque sujet que la stature (x) et la longueur du tibia (y), nous voulons rechercher les représentations graphiques de ces variables propres à nous éclairer à la fois sur la composition du groupe et sur la loi qui lie les deux grandeurs mesurées.

Le tableau de corrélation stature/tibia (fig. 1) centré autour du point moyen C ($X = 174{,}55$; $Y = 39{,}43$) constitue une première représentation très simple de l'échantillon qui permet déjà de répondre à certaines questions. L'étalement du nuage indique l'importance des déviations standard de chaque variable; son applatissement révèle la force de la corrélation ($\rho = 0{,}783$) et donne une première idée grossière de la loi qui lie la longueur du tibia à la stature.

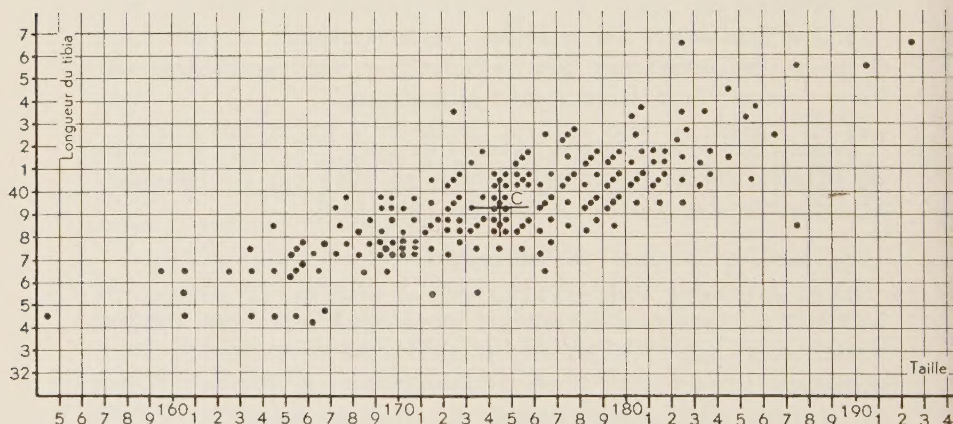


Fig. 1. — Diagramme de dispersion. Echantillon de 192 étudiants de l'U. L. B. C est le point figuratif des moyennes de la taille et du tibia.

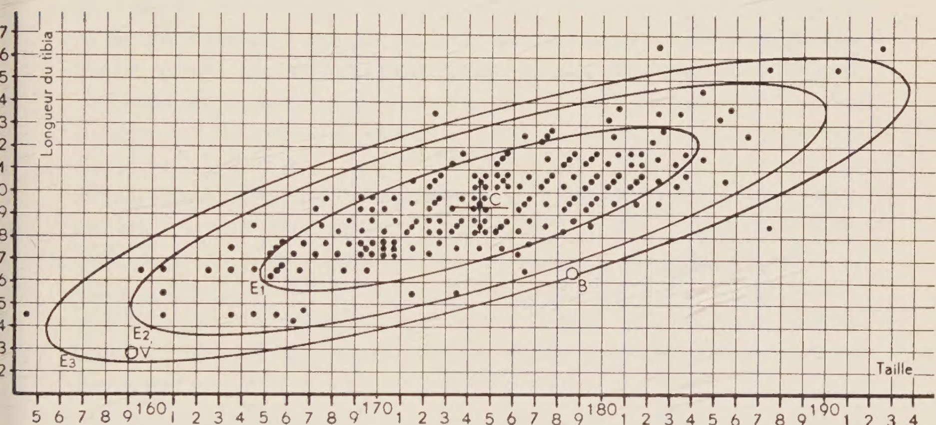
La situation sur la figure 1 d'un point dont les coordonnées sont égales respectivement à la stature et à la longueur du tibia d'un sujet étranger, indique approximativement si ce sujet peut être rangé dans le groupe des étudiants (du moins en ce qui concerne sa taille et son tibia). La réponse est affirmative si le point-image du sujet se place franchement à l'intérieur du nuage. Mais si ce point tombe dans une région où les points

se font plus rares, on désirera, avant de se prononcer, préciser le degré de l'éloignement avec le point moyen.

Cette précision est fournie par les ellipses équiprobables tracées autour du point moyen comme centre (fig. 2). Ces courbes estimées à partir de l'échantillon, se rapportent à la population totale des étudiants.

Remarquons immédiatement que l'emploi des ellipses est subordonné à la condition de normalité de la distribution simultanée de la stature (x) et du tibia (y).

Heureusement, un très grand nombre de répartitions rencontrées en biologie peuvent être considérées comme sensiblement normales, de sorte que cette condition n'est pas très restrictive en pratique.



ig. 2. — E_1, E_2, E_3 sont les trois ellipses équiprobables contenant respectivement 70, 95 et 99 % de la population totale des étudiants, dans l'hypothèse où la répartition bivariée de la taille et du tibia est normale. Les points représentent l'échantillon des 192 sujets. B figure un étudiant hongrois (taux d'éloignement : 98,7 %). V figure un étudiant polonais (taux d'éloignement : 98,4 %).

Après avoir vérifié que notre échantillon appartient effectivement à une population où la stature et le tibia se répartissent normalement (11), nous écrivons l'équation des ellipses équi-

(11) YULE, G. U. et KENDALL, M. G., 1946. On trouvera, p. 232, un test simple pour vérifier si une double série de mensurations appartient à un univers bivarié normal.

probables telle qu'on la trouve dans les traités de statistiques (12)

$$\frac{1}{1-\rho^2}\left\{\frac{(x-X)^2}{\sigma^2_x} - \frac{2\rho\,(x-X)\,(y-Y)}{\sigma_x\,\sigma_y} + \frac{(y-Y)^2}{\sigma^2_y}\right\} = \lambda^2 \quad (1).$$

Les meilleures estimations des paramètres X, Y, σ_x, σ_y, ρ (« population values ») sont précisément les valeurs correspondantes calculées directement à partir de l'échantillon (« sample values ») et qui figurent dans le tableau I (13).

L'équation (1) montre qu'à chaque valeur de λ^2 correspond une ellipse particulière et pour des valeurs croissantes de λ^2, ces ellipses contiennent un pourcentage de plus en plus grand de la population totale des étudiants.

Ce pourcentage est fourni par une table de répartition de la grandeur χ^2, que l'on trouve dans la plupart des traités de statistiques. En effet, λ^2 a, dans le cas de deux variables x et y, la même distribution que χ^2 pour deux degrés de liberté (14).

Tableau II.

Extrait d'une table de χ^2 pour 2 degrés de liberté (car il y a 2 variables).

Pourcentage des étudiants situés en dehors des ellipses	$\chi^2 = \lambda^2$
99	0,02
98	0,04
95	0,10
90	0,21
80	0,45
70	0,71
50	1,39
30	|2,41|
20	3,22
10	4,60
5	|5,99|
2	7,82
1	|9,21|
0,1	13,81

(12) Voir par exemple Cramèr, H., 1946, p. 288.
(13) Kendall, M. G., 1946, p. 37, vol. II.
(14) Cramèr, H., 1946, p. 319.

La table II indique qu'il faut donner successivement à λ^2 ($= \chi^2$) les valeurs 2,41; 5,99; 9,21 pour obtenir les équations des ellipses E_1, E_2, E_3 de la figure 2 contenant respectivement 70 %, 95 % et 99 % des étudiants.

Voici, à titre d'exemple, l'équation numérique de E_1

$$\frac{1}{1 - 0{,}783^2}\left\{\frac{(x - 174{,}55)^2}{6{,}314^2} - \frac{2 \times 0{,}783 \times (x - 174{,}55)(y - 39{,}43)}{6{,}314 \times 2{,}259} + \frac{(y - 39{,}43)^2}{2{,}259^2}\right\} = 2{,}41$$

Si les étudiants figurant sur le graphique 1 étaient très nombreux, c'est-à-dire si l'effectif de l'échantillon augmentait pour se rapprocher de celui de la population totale, on pourrait joindre les centres des carrés égaux contenant un même nombre de points. Chaque ligne fermée ainsi obtenue serait d'autant plus proche d'une ellipse que le nombre d'étudiants mesurés serait plus élevé et les carrés plus petits. Cette circonstance justifie le nom d'ellipses d'égale fréquence ou *ellipses équiprobables*. Bien entendu, dire que tous les points de l'ellipse E_1 de la figure 2 ont la même fréquence signifie en réalité que tous les carrés égaux (très petits) placés sur E_1 contiennent le même nombre de points.

L'égale fréquence de tous les points de l'ellipse E_1 nous a donné l'idée de considérer ces points comme également « éloignés » du centre C et nous a amené à introduire la notion de *taux d'éloignement*.

Nous allons définir et justifier cette notion du point de vue biométrique, expliquer ses avantages et ses propriétés.

Par définition, nous dirons que (fig. 2)

le taux d'éloignement des points de l'ellipse E_1 est 70 %,

le taux d'éloignement des points de l'ellipse E_2 est 95 %,

le taux d'éloignement des points de l'ellipse E_3 est 99 %.

Ces trois ellipses sont dessinées sur le graphique 2 à cause des valeurs standard de leur taux d'éloignement. Elles font partie de l'infinité des ellipses équiprobables que l'on peut

tracer autour du point C et dont chacune correspond à un taux d'éloignement déterminé, croissant pour les ellipses de plus en plus grandes.

Ainsi, à chaque point P du plan correspond un taux d'éloignement, qui est le même pour tous les points de l'ellipse équiprobable passant par P.

Illustrons par deux exemples la notion de taux d'éloignement.

Exemple 1. — Le point B du graphique 2 représente un étudiant hongrois de l'U. L. B., né à Brno :

$$\text{stature} = x = 178{,}6$$
$$\text{tibia} = y = 36{,}5$$

L'ellipse (non dessinée) qui passe par le point B, correspond à une valeur de λ^2 égale à 8,778, que l'on calcule en se rapportant à l'équation (1)

$$\frac{1}{1-0{,}783^2}\left\{\frac{(178{,}6-174{,}55)^2}{6{,}314^2} - \frac{2 \times 0{,}783 \times (178{,}6-174{,}55)\,(36{,}5-39{,}43)}{6{,}314 \times 2{,}259} + \frac{(36{,}5-39{,}43)^2}{2{,}259^2}\right\} = 8{,}778$$

Le tableau II indique que pour $\lambda^2 = \chi^2 = 8{,}778$, la probabilité p est comprise entre 2 % et 1 %. Une interpolation donne p = 1,3 %. Cela signifie que l'ellipse (non dessinée) qui passe par le point B contient 98,7 % des étudiants belges. Autrement dit, sur 1000 étudiants belges de l'Université Libre de Bruxelles, seulement 13 sont plus « éloignés » du point moyen que l'étudiant hongrois.

Suivant nos conventions, le taux d'éloignement du sujet B par rapport au groupe belge est 98,7 %. Ce taux, compris entre 95 et 99 %, indique une différence significative entre l'étranger et les sujets belges.

Exemple 2. — Le point V du graphique 2 représente un étudiant polonais de l'U. L. B., né à Vienne.

$$\text{stature} = x = 159{,}1$$
$$\text{tibia} = y = 32{,}9$$

L'ellipse qui passe par V correspond à une valeur de λ^2 égale à

$$\frac{1}{1-0{,}783^2}\left\{\frac{(159{,}1-174{,}55)^2}{6{,}314} - \frac{2\times 0{,}783\times(159{,}1-174{,}55)(32{,}9-39{,}43)}{6{,}314\times 2{,}259} + \frac{(32{,}9-39{,}43)^2}{2{,}259^2}\right\} = 8{,}442$$

Cette ellipse contient 98,4 % des étudiants belges (voir table de χ^2 au tableau II).

Le taux d'éloignement du sujet V est 98,4 %. Cette valeur indique une différence significative entre le sujet V et le groupe belge.

COMPARAISON DES SUJETS B ET V.

Bien que les mesures du sujet B prises séparément le rapprochent des étudiants belges, la combinaison anormale d'un tibia petit avec une taille assez grande l'éloigne du groupe de façon significative. Au contraire, le sujet V a une stature et un tibia en harmonie avec la structure des Belges, mais il s'éloigne du centre C à cause de la valeur faible de ses deux mensurations. Du point de vue biométrique, V est un peu moins « éloigné » du groupe que le sujet B, car 98,4 % < 98,7 %.

Ces considérations montrent l'utilité de la notion de taux d'éloignement qui tient compte non seulement des valeurs des variables prises séparément, mais aussi de la liaison existant entre ces variables.

Si, en comparant les sujets B et V au groupe belge, on s'était contenté de mesurer sur la fig. 2 les distances géométriques BC et VC, de B et V au centre C du graphique, on serait arrivé à la conclusion, biologiquement fausse, que B est plus proche du groupe belge que le sujet V. Des erreurs de ce genre peuvent être évitées par l'emploi du taux d'éloignement.

Avant de montrer comment les ellipses équiprobables et les taux d'éloignement permettent de poser avec précision et d'élucider certaines questions dans des domaines aussi divers que ceux de la paléontologie, de la croissance, de la systématique humaine, nous esquissons rapidement quelques propriétés de ces ellipses.

2. — Propriétés des ellipses équiprobables.

2.1 Caractères biométriques usuels.

La famille des ellipses équiprobables figure de façon complète la distribution simultanée des deux mensurations au sein de la population totale supposée normale. Nous allons montrer comment, à partir du tracé d'une seule ellipse équiprobable correspondant à un taux d'éloignement donné, on peut retrouver graphiquement les caractères biométriques usuels : moyennes, déviations standard, coefficient de corrélation et droites de regression (fig. 3).

Les moyennes. — Les coordonnées du point C, centre de l'ellipse, lues sur les axes, sont les moyennes X et Y.

Les déviations standard. — Les tangentes t_1 et t_2, parallèles à l'axe des y, coupent l'axe des x en deux points de coordonnées $X - \lambda\sigma_x$, $X + \lambda\sigma_x$. Il suffit donc de tracer l'une de ces tangentes pour trouver immédiatement la valeur de $\lambda\sigma_x$.

L'ellipse E_2 (fig. 3) contient 95 % des étudiants, donc $\lambda^2 = 5{,}99$ et $\lambda = \sqrt{5{,}99} = 2{,}448$.

En divisant $\lambda\sigma_x$ par 2,448, on trouve σ_x.

De la même façon, on détermine $\lambda\sigma_y$ (donc σ_y) en menant la tangente t_3 parallèle à l'axe des x jusqu'à son intersection avec l'axe y.

Le coefficient de corrélation. — La parallèle (*a*) menée à l'axe des x par le point C coupe l'ellipse E_1 en un point A dont l'abscisse vaut $X + \lambda\sigma_x \sqrt{1 - \rho^2}$. En divisant le segment CA par $\lambda\sigma_x$, on obtient $\sqrt{1 - \rho^2}$ d'où l'on tire la valeur du coefficient de corrélation ρ.

Droite de régression de y en x. — En joignant les points de contact R_1 et R_2 des tangentes t_1 et t_2, on obtient la droite de régression de y par rapport à x (fig. 3). L'équation de cette droite est

$$y^\star - Y = \rho \frac{\sigma_y}{\sigma_x} (x - X)$$

c'est-à-dire $y^\star - 39{,}43 = 0{,}280\ (x - 174{,}55)$. Elle permet de trouver quelle valeur moyenne $y^\star$ du tibia correspond à une valeur donnée de la stature x; autrement dit, $y^\star$ est la moyenne des tibias de tous les étudiants du groupe qui ont une même taille égale à x.

Droite de régression de x en y. — C'est la droite R_3R_4 qui joint les points de contact des tangentes t_3 et t_4 (fig. 3). Son équation est

$$y - Y = \frac{\sigma_y}{\rho\sigma_x} (x^\star - \chi)$$

c'est-à-dire $y - 39{,}43 = 0{,}457 (x^\star - 174{,}55)$. Cette droite permet d'évaluer quelle moyenne $x^\star$ de la taille correspond à une longueur donnée du tibia (y). On utilise notamment une telle droite de régression lorsqu'on veut se faire une idée de la stature d'un squelette incomplet dont on possède le tibia.

2.2 Loi approchée entre stature et tibia.

Lorsque le coefficient de corrélation est élevé, l'ellipse E_2 (fig. 2) et le nuage de points qu'elle contient sont très allongés, la relation entre x et y se rapproche alors d'une liaison fonctionnelle (où à une valeur de x correspond une seule valeur de y, et vice versa) que l'on peut tacher de figurer par une droite.

La méthode des moindres carrés fournit une solution à ce problème délicat : le grand axe de l'ellipse E_2 étant tel que la somme des carrés des distances de tous les points du groupe à l'axe est minimum (15), K. Pearson propose de l'adopter pour représenter la loi moyenne qui lie la stature et le tibia (16).

Mais alors que les droites de régression fournissent une réponse précise à la question : quelle est la moyenne des y (ou des x) de tous les sujets qui ont un même x (ou y), la droite placée sur le grand axe de l'ellipse ne donne qu'une liaison approchée entre x et y; en réalité, le lien entre x et y est de nature stochastique (à chaque valeur de x correspondent plusieurs valeurs de y, et inversement) et ne peut pas avec rigueur être remplacé par une liaison fonctionnelle. Cependant, en pratique, il est parfois avantageux — notamment pour comparer l'allure de la croissance dans des populations différentes (17) — de représenter globalement le nuage de points par le grand axe de l'ellipse E_2 (qui se superpose d'ailleurs aux grands axes de toutes les ellipses équiprobables).

(15) Cramèr, H., 1946, p. 275.
(16) Pearson, K., 1901, p. 559.
(17) Twiesselmann, F., 1949, p. 70.

2,3 INVARIANCE.

Dans ce qui précède, les deux mesures x et y étaient exprimées avec une même unité (le cm) et, sur les graphiques, une même unité était choisie sur les deux axes x et y. Une même longueur représentait donc 1 cm de la taille sur l'axe des x et 1 cm du tibia sur l'axe des y, ce qui était naturel parce que les deux mensurations envisagées ont des ordres de grandeur comparables.

Cependant, la plupart des notions et propriétés relatives aux ellipses équiprobables (notamment celles du § 2.1) sont indépendantes de ces restrictions : elles restent valables si les deux mesures envisagées x, y, ne sont pas exprimées avec une même unité et si les unités choisies pour graduer les deux axes x et y du graphique ont des longueurs différentes (voir fig. 7).

En d'autres termes, ces propriétés sont *invariantes* vis-à-vis de tout changement d'unités, c'est-à-dire vis-à-vis des changements de variables du type $x' = \alpha x$, $y' = \beta y$, où α et β sont deux nombres quelconques.

Lorsqu'on change d'unités de mesure ou qu'on modifie la graduation des axes x et y, la forme des ellipses équiprobables est, en général, altérée.

Cependant, la valeur du coefficient de corrélation reste inchangée.

De plus, un même sujet aura toujours un même taux d'éloignement, un même λ^2.

Le taux d'éloignement d'un sujet, ainsi que son λ^2, tout comme le coefficient de corrélation, ne dépendent ni des unités de mesure, ni des unités choisies sur les axes de la représentation graphique. Ce sont des *invariants* vis-à-vis de tout changement d'unités.

Ainsi, que la stature et le tibia soient évalués en pouces ou en cm, et quelles que soient les longueurs respectives des unités sur les axes, on trouve toujours le même coefficient de corrélation pour un même groupe, et le taux d'éloignement d'un sujet est toujours le même par rapport à un groupe donné.

De même, *les procédés graphiques décrits au § 2.1 pour déduire d'une ellipse équiprobable les caractères biométriques usuels (moyennes, déviations standard, coefficient de corrélation, droites de régression) restent inchangés, quelles que soient les unités.*

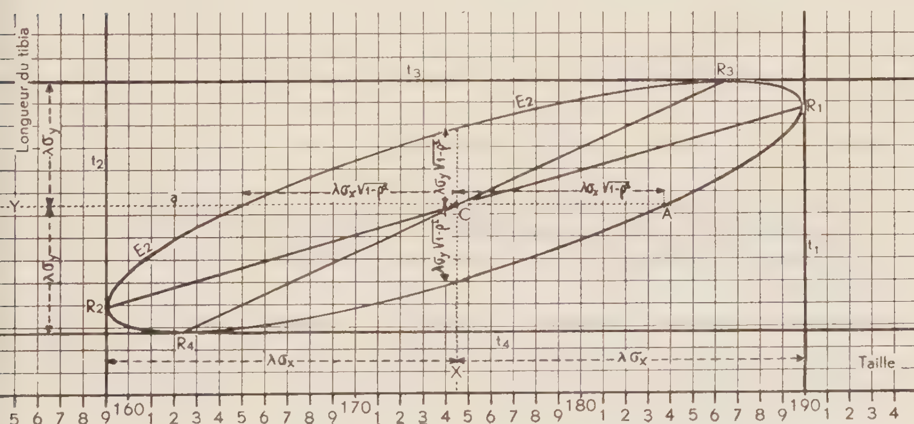
Ainsi donc, pour les propriétés invariantes que nous venons d'énumérer, x et y peuvent désigner des mesures exprimées en unités quelconques et être reportés sur un graphique où les unités

sur les deux axes sont égales ou non. Cette invariance sera effectivement mise à profit si l'on envisage deux mensurations x, y, avec des ordres de grandeur très différents, comme la stature et la hauteur du nez : dans un tel cas, en effet, une même longueur sur les deux axes x, y, représentera en général deux grandeurs différentes (par exemple, 1 cm pour la stature et 1 mm pour la hauteur du nez).

Ces propriétés invariantes pourront également être utilisées lorsque les grandeurs mesurées sont de nature différente (par exemple, une longueur et un poids). Notons cependant que, dans ce cas, les conditions initiales d'emploi des ellipses équiprobables sont moins souvent satisfaites. Ainsi, on ne peut pas circonscrire rigoureusement par des ellipses les nuages de points figurant le poids et la stature d'un groupe : cela provient de ce que la distribution simultanée poids-taille n'est pas normale.

L'invariance est un caractère essentiel pour des grandeurs chargées d'exprimer une réalité biologique. C'est une des raisons qui nous portent à croire que le taux d'éloignement peut rendre, en biologie, des services analogues à ceux du coefficient de corrélation.

Ajoutons une dernière remarque. Contrairement aux notions citées ci-dessus, les axes des ellipses équiprobables ne jouissent



. 3. — L'ellipse E_2 contient 95 % de la population totale des étudiants. t_1, t_2; t_3, t_4 sont les tangentes à l'ellipse parallèles aux axes. R_1R_2 est la droite de régression de y en x. R_3R_4 celle de x en y. La différence entre l'abscisse de t_1 et celle de C est $\lambda\sigma_x$ (λ vaut ici $\sqrt{5{,}99} = 2{,}448$). La distance CA vaut $\lambda\sigma_x\sqrt{1-\rho^2}$. Le grand axe de l'ellipse n'est pas tracé parce qu'il est très proche de R_1R_2.

pas du même caractère d'invariance. En d'autres termes, ces axes ne se correspondent pas dans deux ellipses dont l'une est transformée de l'autre par un changement d'unités. Un sujet placé sur une extrémité du grand axe de l'ellipse E_2 (fig. 3) ne se trouvera pas en général sur le grand axe de l'ellipse, transformée de E_2, construite sur un autre graphique dont les axes x et y sont gradués différemment. Cela montre que le grand axe de E_2 (fig. 3) n'est pas la seule solution au problème consistant à figurer par une relation fonctionnelle le lien entre x et y. Selon la représentation graphique choisie, on aura chaque fois des droites différentes pour figurer le lien entre x et y (toutes ces droites sont situées entre les droites de régression). C'est pour éviter cet inconvénient, que JONES (1937), TEISSIER (1948), KERMACK et HALDANE (1950) (18) suggèrent de choisir parmi toutes les droites, le grand axe de la distribution des deux variables exprimées en unités standard $\frac{x}{\sigma_x}$ et $\frac{y}{\sigma_y}$. Cette droite est l'une des diagonales (non tracées) du rectangle circonscrit à l'ellipse de la figure 3; elle a comme pente $-\frac{\sigma_y}{\sigma_x}$ et est invariante pour les changements d'échelle. La somme des produits des distances prises parallèlement aux axes, de chaque point figuratif à cette droite, est minimum.

3. — GÉNÉRALISATION DU TAUX D'ÉLOIGNEMENT LORSQU'IL Y A PLUS DE DEUX VARIABLES.

Avant de donner quelques exemples d'utilisation des ellipses équiprobables, nous allons montrer comment la notion de taux d'éloignement peut être généralisée au cas de plusieurs variables.

La description d'une population est évidemment d'autant plus précise que l'on dispose de plus de caractères mesurables. C'est uniquement parce que la méthode si pratique des nuages de points et des ellipses est réalisable graphiquement, que nous nous sommes limités jusqu'à présent à deux variables. Il arrive d'ailleurs que, sans perdre beaucoup d'information, on puisse éliminer certaines mensurations lorsqu'elles présentent avec les autres de fortes corrélations. Parfois aussi, il est possible

(18) JONES, H. E. (1937), p. 21;
TEISSIER, G. (1948), p. 28;
KERMACK, K. A. et HALDANE, J. B. S. (1950), p. 30.

de grouper les mensurations en formant des fonctions simples, comme des indices $\left(\frac{x}{y}\right)$, des sommes $(x + y)$, etc., et de réduire ainsi à deux le nombre primitif de variables.

Mais souvent le nombre de variables indispensables est supérieur à deux et la représentation graphique n'est plus possible que partiellement.

On peut alors définir, relativement à l'ensemble des variables, un taux d'éloignement qui mesurera d'une façon très fine la situation d'un sujet par rapport au groupe. On se souvient que le sujet B (p. 8) n'était aberrant ni pour la taille ni pour le tibia, mais son taux d'éloignement relatif aux deux mensurations (98,7 %) le rejettait hors du groupe. De même, le taux d'éloignement relatif à trois variables peut indiquer qu'un sujet est aberrant pour un groupe, alors que les trois taux d'éloignement de ce sujet pour les variables prises deux à deux le situent chaque fois à l'intérieur du groupe.

Un taux d'éloignement est donc d'autant plus sensible qu'il se rapporte à plus de variables.

Voici sa définition pour p variables.

Soient $x_1, x_2, \ldots x_p$ les écarts de chaque variable à sa moyenne; $\sigma_1, \sigma_2, \ldots, \sigma_p$ les déviations standard; ρ_{ij} le coefficient de corrélation des variables i et j. Supposons de plus que la distribution simultanée des p variables est normale.

Formons le déterminant des coefficients de corrélation

$$\omega = \begin{vmatrix} 1 & \rho_{12} & \rho_{13} & \ldots & \rho_{1p} \\ \rho_{12} & 1 & \rho_{23} & \ldots & \rho_{2p} \\ \cdot & \cdot & \cdot & \ldots & \cdot \\ \rho_{1p} & \rho_{2p} & \rho_{3p} & \ldots & 1 \end{vmatrix}$$

et appelons ω_{ij} le mineur de la i^{me} rangée et de la j^{me} colonne.

Dans ces notations, la distribution des p variables est

$$dF = \frac{1}{(2\pi)^{\frac{p}{2}}\, \omega^{\frac{1}{2}}}\, e^{-\frac{\lambda^2}{2}}\, \frac{dx_1}{\sigma_1} \cdot \frac{dx_2}{\sigma_2} \cdots \frac{dx_p}{\sigma_p}$$

avec
$$\lambda^2 = \frac{1}{\omega} \Sigma\, \omega_{ij} \frac{x_i\, x_j}{\sigma_i\, \sigma_j}.$$

A chaque valeur de λ^2 correspond un ellipsoïde équiprobable contenant chaque fois un pourcentage donné des sujets du groupe. A un sujet, dont les écarts aux moyennes du groupe sont d_1, d_2, ..., d_p, correspond une valeur de λ^2 égale à

$$\lambda^2 = \frac{1}{\omega} \Sigma\, \omega_{ij} \frac{d_i\, d_j}{\sigma_i\, \sigma_j}$$

λ^2 est distribué comme χ^2 avec p degrés de liberté (19). Une table de χ^2 indique combien de sujets du groupe sont moins aberrants que le sujet donné.

Ce pourcentage est le taux d'éloignement du sujet.

Par exemple, dans le cas de trois variables, on a

$$\lambda^2 (1 - \rho^2_{12} - \rho^2_{13} - \rho^2_{23} + 2\rho_{12}\rho_{13}\, \rho_{23}) = (1 - \rho^2_{23}) \left(\frac{d_1}{\sigma_1}\right)^2 + (1 - \rho^2_{13}) \left(\frac{d_2}{\sigma_2}\right)^2 + (1 - \rho^2_{12}) \left(\frac{d_3}{\sigma_3}\right)^2 + 2\,(\rho_{23}\rho_{13} - \rho_{12}) \frac{d_1 d_2}{\sigma_1\, \sigma_2} + 2\,(\rho_{12}\rho_{23} - \rho_{13}) \frac{d_1 d_3}{\sigma_1\, \sigma_3} + 2\,(\rho_{13}\rho_{12} - \rho_{23}) \frac{d_2 d_3}{\sigma_2\, \sigma_3}$$

Les valeurs de λ^2 qui correspondent aux taux d'éloignement de 70 %, 95 % et 99 % sont respectivement, pour 3 degrés de liberté, 3,665; 7,815; 11,341.

Ici encore, comme dans le cas de deux variables, λ^2, le taux d'éloignement et les coefficients de corrélation ρ_{ij} sont des invariants pour les changements d'échelle dans les mensurations.

4. — Exemples d'utilisation pratique des ellipses équiprobables et des taux d'éloignement.

4.1 En systématique animale.

Prenons comme premier exemple une population, ou plutôt une thanatocénose (20) de 194 spécimens du genre *Carbonicola* de la famille des *Anthracosiidæ* (21), récoltés par A. Pastiels

(19) Cramèr, H., 1946, p. 319.
(20) Pastiels, A., 1953.
(21) Famille de Mollusques d'eau douce du Carbonifère supérieur.

dans le bassin houiller de Charleroi. Sur les figures 4 et 5, les nuages de points représentent ces *Carbonicola;* ils sont entourés par des ellipses équiprobables contenant 70 et 95 % des spécimens. Les variables sont les indices formés par certains diamètres caractéristiques.

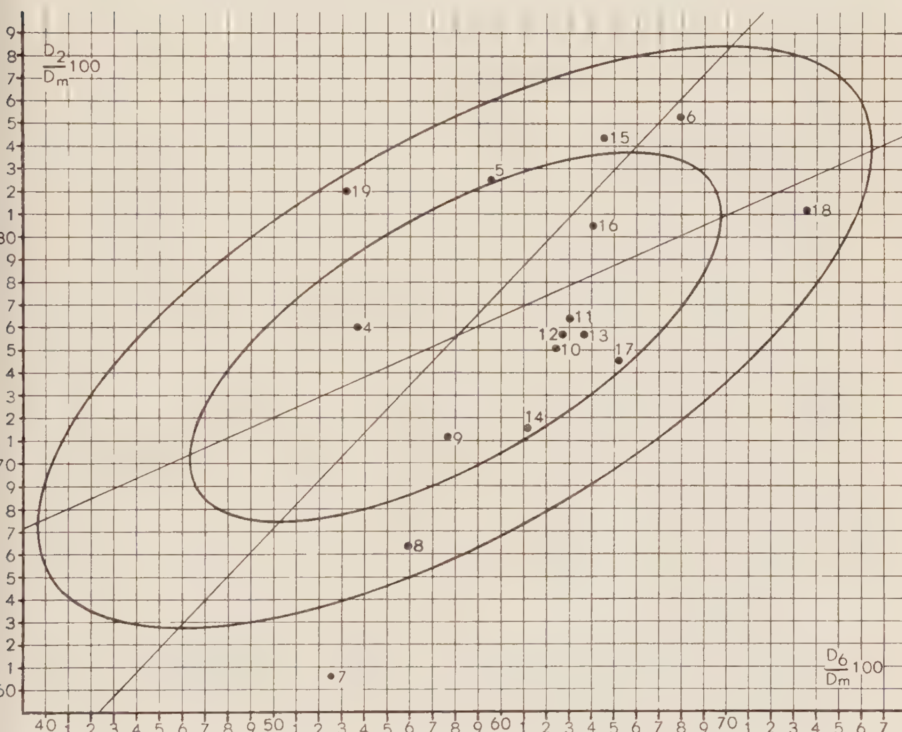


Fig. 4. — Ellipses contenant respectivement 70 et 95 % d'une population de 194 *Carbonicola* de la famille des *Anthracosiidæ*. D_m est le diamètre maximum joignant l'umbo au rebord libre. D_2, D_4, D_6 sont des diamètres radiants de l'umbo et formant avec D_m des angles respectivement égaux à 30°, 60° et 90°. Les points, accompagnés d'un chiffre, représentent des espèces voisines du groupe de *C. Communis* déterminées anciennement. Ainsi 4 est le type de *C. pyramidata,* 5 celui de *C. obliquissima,* 7 celui de *C. browni,* etc (voir PASTIELS, A., p. 35). Dessin repris du travail de M. PASTIELS avec l'autorisation de l'Assoc. pour l'Etude de la Paléont. et de la Strat. Houillères.

Les types d'espèces de *Carbonicola* déterminés anciennement et se rapprochant du groupe de A. PASTIELS sont aussi placés sur ces graphiques.

Rappelons quelques conclusions de l'auteur :

1. Les espèces du genre *Carbonicola* ne peuvent être caractérisées que par des variables continues.

2. On n'a pas pu montrer — malgré de très nombreux essais portant sur toutes les mensurations et leurs indices — que le groupe récolté est mixte. Il faut donc admettre à preuve du contraire qu'il est homogène.

3. Aucun des types taxonomiques anciens ne groupe autour de lui des spécimens récoltés par l'auteur (voir figures 4 et 5).

L'ensemble des types taxonomiques voisins paraît se superposer au groupe actuel.

Sans vouloir entrer dans les discussions délicates sur la nomenclature et passer en revue toutes les espèces voisines, on voit cependant que le type taxonomique numéroté 7 par l'auteur a, sur chaque graphique, un taux d'éloignement supérieur à 95 %. Il s'agit donc là, sans doute, d'une espèce réellement distincte du groupe de l'auteur.

Au contraire, les types taxonomiques n° 11, 12, 13, par exemple, sont bien plus proches entre eux sur chaque graphique que ne le sont trois spécimens quelconques du groupe de l'auteur. Il est difficile par conséquent de les élever au rang de types taxonomiques distincts.

Nous croyons que l'embarras où se trouve le systématicien qui opère sur une population nombreuse et homogène, alors que ses prédécesseurs ont défini des types taxonomiques voisins à partir de spécimens uniques ou peu nombreux, sera un peu diminué s'il définit sa population par un type proche du type moyen et s'il fixe comme limites à sa population des ellipses équiprobables ainsi que l'a fait A. PASTIELS.

Il nous semble aussi que le taux d'éloignement (pour plusieurs variables) de chacun des types taxonomiques anciens par rapport au groupe actuel, peut servir de base au reclassement de toutes les espèces voisines.

Si ce taux était supérieur à 70 % ou à 95 % (les systématiciens auraient à fixer eux-mêmes la limite), on conviendrait de considérer le type ancien distinct du groupe actuel. Si ce taux était inférieur à ces limites, on devrait admettre que la morphologie

du type taxonomique ancien ne se distingue en rien de celle de la population actuelle.

Nous croyons que le cas des *Carbonicola* de A. PASTIELS n'est pas unique et qu'il arrive fréquemment qu'une population homogène porte en bordure les types taxonomiques voisins, comme si l'on avait anciennement choisi comme types ce qui n'étaient que les spécimens les plus aberrants dans diverses directions d'une seule et même espèce.

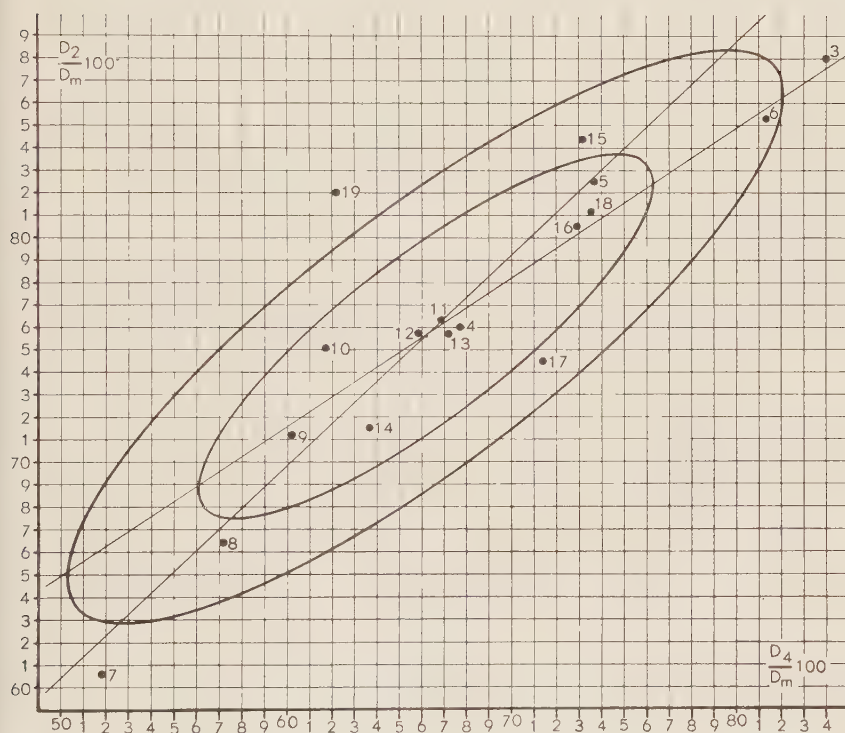


Fig. 5. — Ellipses contenant respectivement 70 et 95 % d'une population de 194 *Carbonicola* de la famille des *Anthracosiidæ*.

4.2 EN SYSTÉMATIQUE HUMAINE.

A l'occasion de son étude anthropologique des habitants de l'Oasis de Kharga (22), F. TWIESSELMANN confronta cette popu-

(22) TWIESSELMANN, F., 1951.

lation méditerranéenne, d'une part avec des Ethiopiens, représentés par les Galla, d'autre part avec des Nilotes, représentés par des Chillouk, des Nouer, des Anouak et des Nouba.

Sur le graphique (fig. 6) indiquant la taille symphysaire et la largeur du bassin, les Khargiens et les Nilotes se groupèrent

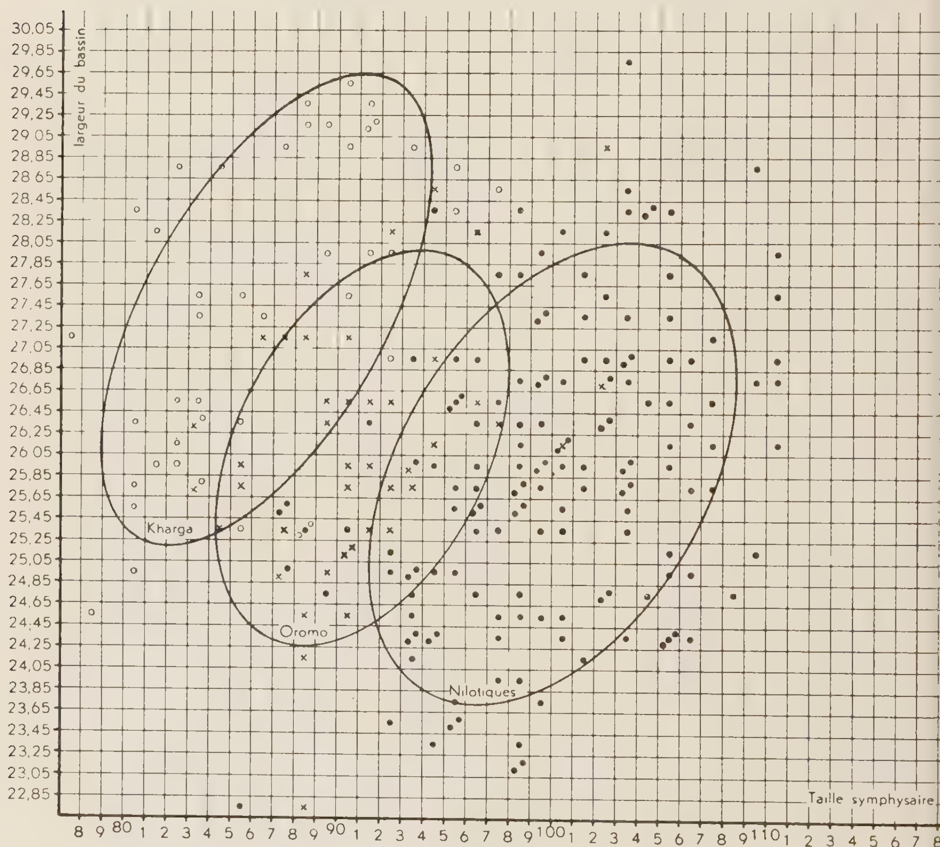


Fig. 6. — Echantillons de 52 Khargiens, 42 Oromo, 150 Nilotes, chacun entouré d'u ellipse contenant 70 % des sujets. Les effectifs de 52 et 42 sujets suffisent pour u comparaison entre des populations. Bien entendu, si l'on voulait donner une descript détaillée de chaque groupe, il faudrait estimer les paramètres des ellipses à partir d'effec supérieurs de préférence à 100 sujets. On voit que le glissement d'une population l'autre ne se fait pas dans le sens du grand axe qui est pratiquement celui de la cro sance, mais plutôt latéralement, la distinction entre les populations étant due surt aux différences entre les proportions. Sans ellipses, les nuages de points seuls n'aurai pas donné une vision aussi claire des situations relatives des trois populations.

en deux nuages de points distincts, recouverts chacun en partie par les points représentatifs des Gallas. F. TWIESSELMANNN nous suggéra de compléter ce graphique par trois ellipses équiprobables contenant chacune 70 % des sujets de l'une des populations (fig. 6).

Les situations respectives des trois groupes sont ainsi précisées; les trois ellipses nettement distinctes ont des axes sensiblement parallèles : cela indique que les accroissements de la largeur du bassin et de la taille symphysaire sont dans un même rapport chez les trois populations. Les Oromo-Galla occupent une position intermédiaire entre les Khargiens et les Nilotes, comme ils le font d'ailleurs pour d'autres caractères anthropologiques, tels la pigmentation et l'indice céphalique, les caractères descriptifs de la face, etc...

La variation d'un groupe à l'autre est certes due à des différences entre les moyennes de la taille symphysaire et la largeur du bassin; mais c'est essentiellement par les proportions du corps (et c'est là ce qui intéresse surtout l'anthropologiste) que les trois populations se distinguent : le glissement de l'une à l'autre ne se fait pas dans le sens du grand axe des ellipses, mais plutôt latéralement; les Nilotes ont, malgré de plus longues jambes, un bassin plus étroit que les deux autres groupes.

Ainsi, les ellipses équiprobables fournissent une base concrète aux discussions concernant les différences morphologiques entre ces trois populations.

Une telle représentation graphique est encore indispensable à celui qui cherche à expliquer les différences entre les populations, soit par l'effet de métissages successifs, soit par l'adaptation à des milieux très différents.

La première hypothèse — qui n'est jamais vérifiable par les seules méthodes de l'anthropologie physique et de la génétique, empressons-nous de le dire — ne sera en tout cas acceptée que s'il est possible de placer entre deux populations supposées primitives, le groupe que l'on croit résulter de leur croisement. Si le métissage est récent, le groupe intermédiaire devra avoir une variabilité supérieure à celle des populations d'origine.

La deuxième hypothèse, que F. TWIESSELMANN juge digne d'être approfondie, serait vérifiable dans une certaine mesure si l'on réussissait à établir une corrélation entre l'habitat et la morphologie des habitants. Il se pourrait, en effet, que placés sur des graphiques taille symphysaire/largeur bassin, de nom-

breuses populations de régions marécageuses auraient, comme les Nilotes, de longues jambes et un bassin étroit par rapport à leurs voisins.

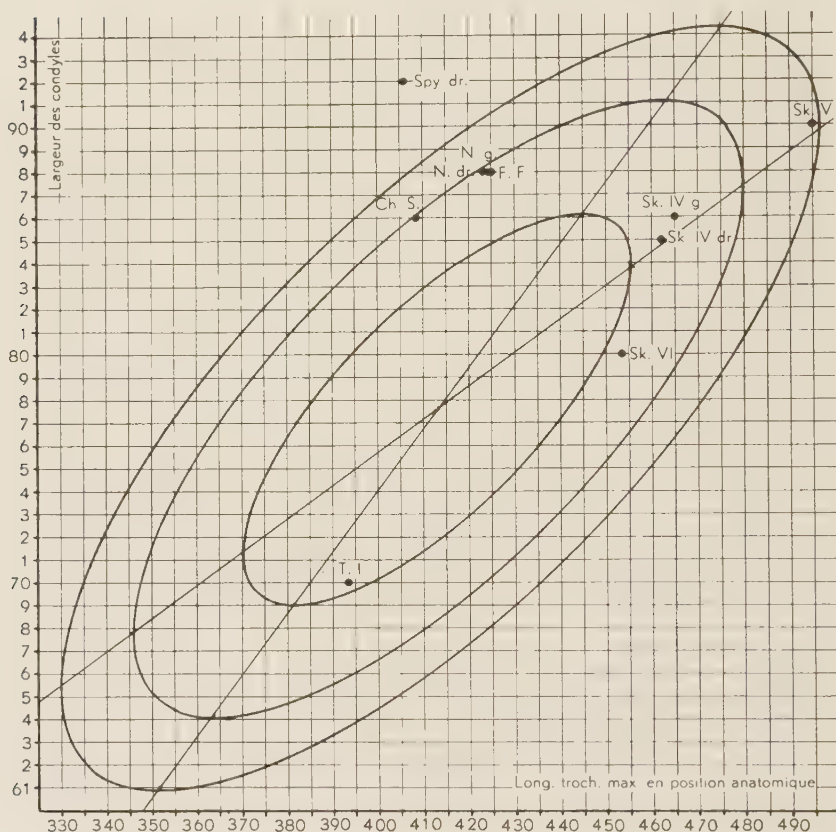


Fig. 7. — Ellipses de 70, 95 et 99 % estimées à partir d'une population de 416 fémurs actuels, droits et gauches, provenant de divers cimetières belges. Les fémurs néanderthaliens d'Europe, Spy, Neanderthal gauche et droit, La Chapelle aux Saints, Fond-de-Forêt, se groupent avec des taux d'éloignement proches de 95 % et dépassant largement 99 % pour celui de Spy; ils sont plus robustes que les fémurs actuels. Les fémurs Skhùl ont des taux d'éloignement supérieurs à 70 %; ils se groupent aussi et leur situation indique qu'ils sont plus longs mais proportionnellement moins robustes que ceux de Neanderthal. Le fémur de la femme de Tabùn (T. 1) est court, mais on peut admettre que si sa croissance avait continué, il aurait glissé parallèlement au grand axe (non tracé) des ellipses pour venir se placer parmi les fémurs Skhùl.

4.3 En paléontologie humaine.

Pour décrire de façon objective des ossements paléolithiques, F. Twiesselmann juge souhaitable de les comparer à une population homogène d'ossements récents, provenant de préférence de la même région que les os anciens. L'ensemble moderne, représenté par des ellipses équiprobables relatives chaque fois à deux variables (mensurations ou indices), contient une large gamme de variabilité et permet de préciser le degré d'aberration des fossiles.

F. Twiesselmann (23) a fait réaliser de tels graphiques pour confronter les fémurs néanderthaliens avec un ensemble homogène de 416 fémurs actuels provenant de divers cimetières belges. Les graphiques 7 et 8 montrent des ellipses équiprobables contenant respectivement 70, 95 et 99 % des fémurs actuels des deux sexes, tant droits que gauches; on a évité une séparation des deux sexes, qui comporte toujours un grand nombre de cas douteux lorsque les squelettes ne sont pas identifiés. Les points représentatifs de fémurs fossiles sont reportés sur les deux graphiques. Les fémurs néanderthaliens d'Europe : Spy, Néanderthal gauche et droit, La-Chapelle-aux-Saints, Fond-de-Forêt, se groupent remarquablement sur les deux graphiques. Sur la figure 7, qui donne la longueur trochantérienne maximum en position anatomique avec la largeur des condyles, les taux d'éloignement sont près d'atteindre 95 % et le fémur de Spy dépasse largement la valeur de 99 %. Ce n'est donc qu'exceptionnellement que les fémurs belges actuels sont aussi robustes que ces cinq fémurs fossiles.

Sur la figure 8, donnant la longueur maximum en position anatomique avec le diamètre transversal au milieu de la diaphyse, les cinq fémurs néanderthaliens d'Europe se groupent sans se distinguer du noyau central de la population actuelle; cependant, tous les cinq ont un diamètre légèrement plus fort que les fémurs actuels de même longueur.

L'usage des ellipses permet ici d'opérer un choix dans les mensurations : pour décrire le caractère paléolithique des fémurs de Néanderthal, on utilisera de préférence les variables qui donnent des taux d'éloignement élevés avec la population actuelle.

Sur le graphique 7, nous avons aussi reporté les fémurs palestiniens de Skhùl et de Tabùn. Les fémurs Skhùl se groupent avec des taux d'éloignement supérieur à 70 %, ils sont plus

(23) Twiesselmann, F., 1954.

longs mais proportionnellement un peu moins robustes que les Néanderthaliens d'Europe et que les fémurs actuels.

Le fémur de la femme de Tabùn est petit et est proche de la limite de l'ellipse de 70 %, mais sa structure ne semble pas essentiellement différente de celle des fémurs Skhùl. Si sa croissance avait continué, le point représentatif de ce fémur aurait sans doute glissé parallèlement au grand axe des ellipses pour venir se placer parmi les fémurs Skhùl.

Sur le canevas des fémurs actuels, les fémurs néanderthaliens d'Europe et de Palestine occupent deux régions distinctes séparées par le grand axe des ellipses équiprobables : le premier groupe est plus robuste, le second plus faible que la population actuelle.

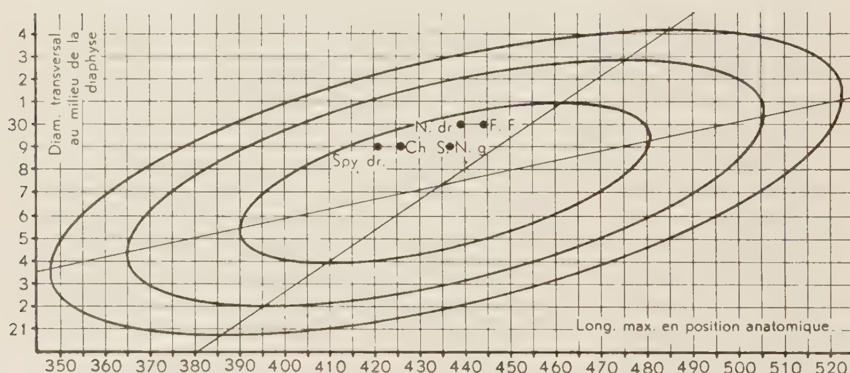


Fig. 8. — Ellipses de 70, 95 et 99 % estimées à partir de 416 fémurs actuels. Les fémurs néanderthaliens se groupent, mais avec des taux d'éloignement inférieurs à 70 %; ils ne se distinguent pas de façon significative de la population actuelle. L'usage des ellipses permet ainsi d'opérer un choix dans les mensurations; pour décrire le caractère paléolithique des fémurs, on utilisera les variables qui donnent des taux d'éloignement élevés avec la population actuelle.

5. — Tracé des ellipses équiprobables.

Pour terminer, nous indiquons brièvement les calculs indispensables au tracé des ellipses à l'aide de l'ellipsographe Riefler. Il s'agit de déterminer les extrémités des axes. Pour fixer les idées, nous supposons que les deux mensurations sont exprimées

en centimètres (c'est-à-dire en mêmes unités). Il faut alors distinguer deux cas suivant que les graduations des axes x et y sont égales ou non.

5.1 Les unités sur les axes sont égales.

Premier cas. — Montrons par exemple comment on a tracé l'ellipse E_2 de la fig. 2 qui contient 95 % des sujets. Sur ce graphique, 1 cm de la taille est représenté sur l'axe des x par un segment égal à celui qui représente 1 cm du tibia sur l'axe des y.

On calcule successivement, suivant les méthodes habituelles

$$\begin{aligned} X &= 174,55 \\ Y &= 39,43 \\ \sigma_x &= 6,314 \\ \sigma_y &= 2,259 \\ \rho &= 0,783 \end{aligned}$$

Ensuite,

$$t = \frac{2\,\rho\ \sigma_x\ \sigma_y}{\sigma^2_x - \sigma^2_y} = 0,642.523.303.$$

$$\left.\begin{aligned} m &= \frac{-1 + \sqrt{1 + t^2}}{t} = 0,293.573.573. \\ m' &= -\frac{1}{m} = -3,406.301.152. \end{aligned}\right\} \text{coefficients angulaires des axes de l'ellipse } E_2$$

On calcule encore

$$\left\{\begin{aligned} a_1 &= \lambda \sqrt{1 - \rho^2} : \sqrt{\frac{1}{\sigma^2_x} - \frac{2\rho m}{\sigma_x \sigma_y} + \frac{m^2}{\sigma^2_y}} = 15,428.584 \\ b_1 &= a_1 m = 4,529.424 \end{aligned}\right.$$

avec $\lambda = \sqrt{5,99} = 2,448.$

$$\left\{\begin{aligned} a_2 &= \lambda \sqrt{1 - \rho^2} : \sqrt{\frac{1}{\sigma^2_x} - \frac{2\rho m'}{\sigma_x \sigma_y} + \frac{m'^2}{\sigma^2_y}} = 0,931.863 \\ b_2 &= a_2 m' = -3,174.208. \end{aligned}\right.$$

avec $\lambda = 2,448.$

Les extrémités des axes de l'ellipse E_2 ont comme coordonnées

$$\left\{\begin{array}{l} X + a_1 = 189{,}97 \\ Y + b_1 = 43{,}95 \end{array}\right. \text{ et } \left\{\begin{array}{l} X - a_1 = 159{,}12 \\ Y - b_1 = 34{,}90 \end{array}\right. \text{ d' une part}$$

$$\text{et } \left\{\begin{array}{l} X + a_2 = 175{,}48 \\ Y + b_2 = 36{,}25 \end{array}\right. \text{ et } \left\{\begin{array}{l} X - a_2 = 173{,}61 \\ Y - b_2 = 42{,}60 \end{array}\right. \text{ d'autre part.}$$

Ces extrémités connues, on peut utiliser l'ellipsographe Riefler pour tracer l'ellipse.

5.2 Les unités sur les axes sont inégales.

Deuxième cas. — Prenons comme exemple le dessin de l'ellipse E_2 de la figure 7.

Sur ce graphique,

x représente la longueur trochantérienne maximum en position anatomique exprimée en mm,

y représente la largeur des condyles exprimée en mm.

Mais on a représenté 1 unité sur l'axe des x par une longueur 5 fois moindre que l'unité sur l'axe y.

Représentons par ν le rapport des deux segments unitaires.

$$\text{On a donc } \frac{\text{unité sur l'axe } x}{\text{unité sur l'axe } y} = \frac{1}{5} = \nu.$$

On calcule par les méthodes habituelles

$$X = 413{,}21 \qquad Y = 77{,}61 \qquad \rho = 0{,}732$$
$$\sigma_x = 27{,}590 \qquad \sigma_y = 5{,}467$$

Mais dans la suite des calculs, pour tenir compte de la différence des segments unitaires, on opère avec $\nu\sigma_x = \sigma'_x$ au lieu de

σ_x, c'est-à-dire avec $\frac{1}{5} \times 27{,}590 = 5{,}518 = \sigma'_x$.

On calcule alors

$$t = \frac{2\rho\sigma'_x\sigma_y}{\sigma'^2_x - \sigma^2_y} = 78{,}831.830.185.$$

$$m = \frac{-1 + \sqrt{1 + t^2}}{t} = 0{,}987.395.222.$$

$$m' = -\frac{1}{m} = -1{,}012.765.686.$$

$$\left\{ \begin{array}{l} a_1 = \lambda \sqrt{1 - \rho^2} : \sqrt{\dfrac{1}{\sigma'^2_x} - \dfrac{2\rho m}{\sigma'_x \sigma_y} + \dfrac{m'^2}{\sigma^2_y}} = 12{,}591.734 \\ b_1 = a_1 m = 12{,}433.018 \qquad (\lambda = 2{,}448) \end{array} \right.$$

$$\left\{ \begin{array}{l} a_2 = \lambda \sqrt{1 - \rho^2} : \sqrt{\dfrac{1}{\sigma'^2_x} - \dfrac{2\rho m'}{\sigma'_x \sigma_y} + \dfrac{m'^2}{\sigma^2_y}} = 4{,}890.400 \\ b^2 = a_2 m' = -4{,}952.830 \qquad (\lambda = 2{,}448) \end{array} \right.$$

Alors les coordonnées des extrémités des axes sont

$$\left\{ \begin{array}{l} X + \dfrac{a_1}{v} = 476{,}16 \\ Y + b_1 = 90{,}05 \end{array} \right. \text{ et } \left\{ \begin{array}{l} X - \dfrac{a_1}{v} = 350{,}25 \\ Y - b_1 = 65{,}16 \end{array} \right. \text{ d'une part}$$

$$\text{et } \left\{ \begin{array}{l} X + \dfrac{a_2}{v} = 437{,}97 \\ Y + b_2 = 72{,}65 \end{array} \right. \text{ et } \left\{ \begin{array}{l} X - \dfrac{a^2}{v} = 388{,}45 \\ Y - b_2 = 82{,}56 \end{array} \right. \text{ d'autre part.}$$

Résumé.

Tout échantillon tiré d'une population et caractérisé par deux mensurations peut être représenté par un diagramme de dispersion. Si l'on admet l'hypothèse que les mensurations ont une distribution simultanée normale, la population elle-même est figurée complètement par une famille d'ellipses équiprobables contenant un pourcentage de plus en plus grand de sujets.

A partir du tracé d'une seule de ces ellipses, on retrouve graphiquement tous les caractères biométriques habituels, moyennes, déviations standard, coefficients de corrélation, droites

de régression, de sorte qu'une seule ellipse synthétise la population, sans que l'on perde l'information donnée primitivement par les paramètres habituels. L'axe de l'ellipse figure approximativement la loi entre les deux mensurations.

Pour concrétiser le degré d'aberration d'un sujet par rapport au groupe, on introduit une nouvelle notion, celle de taux d'éloignement. Ce taux est exprimé par le pourcentage de la population contenu dans l'ellipse équiprobable qui passe par le point représentatif du sujet. Le taux d'éloignement est donc le même pour tous les points placés sur une même ellipse. Il exprime une sorte d'éloignement « biométrique » du sujet par rapport aux moyennes de la population. De même que le coefficient de corrélation, le taux d'éloignement est une grandeur qui reste constante pour tous les changements des unités de mesure. Cette propriété d'invariance est essentielle pour des grandeurs chargées d'exprimer une réalité biologique.

La notion de taux d'éloignement s'étend au cas de plusieurs variables, mais alors la représentation graphique n'est plus possible que partiellement.

Trois exemples illustrent l'utilisation pratique des ellipses équiprobables et des taux d'éloignement.

1. Un groupe de *Carbonicola* récoltés par M. PASTIELS montre

a) que les ellipses équiprobables donnent une description complète du groupe.

b) que les problèmes de reclassement de types taxonomiques voisins, anciennement déterminés à partir d'un petit nombre ou même d'un seul spécimen, se posent de façon plus précise si l'on place ces types sur le graphique de la population et si l'on détermine pour chacun le taux d'éloignement par rapport aux moyennes. De telles représentations graphiques révèlent souvent qu'une population homogène porte en bordure les types taxonomiques voisins, comme si l'on avait anciennement choisi comme types ce qui n'étaient que les spécimens les plus aberrants dans diverses directions d'une seule et même espèce.

2. Les trois populations : Khargiens, Galla et Nilotes, chacune circonscrite par une ellipse, se comportent pour les deux caractères somatiques, taille symphysaire et largeur du bassin, comme elles le font pour d'autres caractères anthropologiques, les Galla se plaçant entre les Khargiens et les Nilotes. La forme et les emplacements respectifs des trois ellipses synthétisent clairement les situations réciproques des trois populations.

Une telle représentation est avantageuse chaque fois que l'on compare des sous-groupes d'une même espèce parce qu'elle permet d'examiner les hypothèses de métissage et d'adaptation au milieu.

3. Pour définir objectivement le degré d'aberration d'ossements paléolithiques, il est essentiel de les comparer, non pas à un seul mais à un ensemble d'ossements modernes, car seule une population entière contient une large gamme de la variabilité des caractères envisagés. L'emploi des ellipses s'impose ici avec force puisque seules les méthodes biométriques sont utilisables en paléontologie.

F. Twiesselmann a réalisé de tels graphiques pour confronter des fémurs néanderthaliens avec un ensemble homogène de fémurs récents. Les deux graphiques que nous reproduisons montrent que les fémurs de Néanderthal sont aberrants de façon significative, mais seulement pour certains caractères; la méthode des ellipses et du taux d'éloignement permet ici de faire un tri dans les mensurations et de ne retenir que celles qui indiquent une aberration significative des fémurs néanderthaliens avec le groupe moderne.

Nous terminons cette note en indiquant les calculs indispensables au tracé des ellipses équiprobables à l'aide de l'ellipsographe Riefler.

Index bibliographique.

Adam, W., 1952, *Expédition océanographique belge dans les eaux côtières africaines de l'Atlantique Sud (1948-1949). Céphalopodes.* (Mém. de l'Inst. royal des Sc. natur. de Belgique, vol. III, fasc. 3, 142 p., 3 pl. h.-t.)

Cramèr, H., 1946, *Mathematical methods of statistics.* (Princeton University Press, 575 p.)

Davenport, C. B., 1903, *Antero-posterior/Dorso-ventral diameter in Lower Valve of Pecten opercularis.* (Proc. Amer. Ac., vol. 39, p. 149.)

Defrise-Gussenhoven, E., 1954, *Croissance et débilité. Etude biométrique.* (Mém. de l'Inst. royal des Sc. natur. de Belgique, n° 128, 70 p., 20 fig.)

DICKSON, J. D. H., 1886, *Appendix to Family Likeness in Stature*. (Proc. Roy. Soc., vol. 40, p. 63.)

GALTON, F., 1885, *Regression towards Mediocrity in Hereditary Stature*. (Journ. of the Anthrop. Inst. of Great Brit. and Ireland, vol. 15, pp. 246-258.)

GLIBERT, M., 1945, *Faune malacologique du Miocène. Les Pélécypodes*. (Mém. du Musée royal d'Hist. natur. de Belgique, n° 103, 266 p., 12 pl. h.-t.)

— , 1949, *Gastropodes du Miocène moyen du bassin de la Loire*. (Mém. de l'Inst. royal des Sc. natur. de Belgique, 2e série, fasc. 30, 240 p., 12 pl. h.-t.)

JOLIVET, P., 1952, *Notes biométriques sur quelques Chrysomeloidea (Coleoptera)*. (Bull. de l'Inst. royal des Sc. natur. de Belgique, tome XXVIII, n° 44, 8 p.)

JONES, H. E., 1937, *Some geometrical considerations in the general theory of fitting lines and planes*. (Metron, tome 13, p. 21.)

KENDALL, M. G., 1946, *The advanced theory of statistics*. (Griffin, Londres, 2 vol., 457 et 521 p.)

KERMACK, K. A. & HALDANE, J. B. S., 1950, *Organic correlation and allometry*. (Biometrika, vol. 37, pp. 30-41.)

MAYR, E., LINSLEY, E. G. & USINGER, R. L., 1953, *Methods and principles of systematic zoology*. (Mac Graw Hill, New York, Londres, 327 p.)

NOWAKOWSKI, T. K. & PERKAL, V., 1952, *Nouvelles méthodes pour déterminer la dépendance entre la stature, le poids et l'âge de la jeunesse (résumé français)*. (Przeglad antropologiczny, tome XVIII, Poznan, pp. 12-33.)

PASTIELS, A., 1953, *Etude biométrique des Anthracosiidæ du Westphalien A de la Belgique. Les Carbonicola du Toit de la Couche « Huit Paumes »*. (Assoc. pour l'Etude de la Paléont. et de la Stratigr. Houillères, publ. n° 16, 56 p., 20 pl.)

PEARSON, K., 1901, *On lines and planes of closest fit to systems of points in space*. (Phil. Mag., tome 6, pp. 559-572.)

PEARSON, K. & LEE, A., 1903, *On the laws of inheritance in Man. I. Inheritance of physical characters*. (Biometrika, vol. 2, pp. 357-462.)

TEISSIER, G., 1948, *La relation d'allométrie : sa signification statistique et biologique*. (Biometrics, tome 4, pp. 14-48.)

TWIESSELMANN, F., 1949, *Contribution à l'étude de la croissance pubertaire de l'Homme*. (Mém. de l'Inst. royal des Sc. natur. de Belgique, 2e série, fasc. 35, 87 p., 26 fig.)

— , 1951, *Expédition anthropologique du Dr D. J. H. Nyéssen. I. L'Oasis de Kharga*. (Bull. de l'Inst. royal des Sc. natur. de Belgique, tome XXVII, n° 14, 36 p.)

— , 1954, *Propos sur l'Anthropologie*. (Volume jubilaire Victor Van Straelen, tome II, pp. 1065-1098, 7 fig.)

VANDERCAMMEN, A., *Revision des Cyrtospirifères du Frasnien de la Belgique* (à paraître dans les Mémoires de l'Inst. royal des Sc. natur. de Belgique).

YULE, G. U. & KENDALL, M. G., 1946, *An introduction to the theory of statistics*. (Ch. Griffin, Londres, 570 p.)

TABLE DES MATIÈRES.

INSTITUT ROYAL DES SCIENCES NATURELLES DE BELGIQUE.

AD. GOEMAERE, Imprimeur du Roi, 21, rue de la Limite, Bruxelles

AD. GOEMAERE, Imprimeur du Roi, 21, rue de la Limite, Bruxelles